

0590
1026

#2

OIKE

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/970,921

DATE: 10/19/2001

TIME: 10:37:39

Input Set : A:\2428-0108P.txt

Output Set: N:\CRF3\10192001\I970921.raw

ENTERED

3 <110> APPLICANT: Frank Michiels et al.
 5 <120> TITLE OF INVENTION: Improved Barstar Gene
 7 <130> FILE REFERENCE: 2428-0108P
 C--> 9 <140> CURRENT APPLICATION NUMBER: US/09/970,921 OK
 10 <141> CURRENT FILING DATE: 2001-10-05
 12 <160> NUMBER OF SEQ ID NOS: 10
 14 <170> SOFTWARE: PatentIn Ver. 2.0
 16 <210> SEQ ID NO: 1
 17 <211> LENGTH: 270
 18 <212> TYPE: DNA
 19 <213> ORGANISM: Bacillus amyloliquefaciens
 21 <220> FEATURE:
 22 <221> NAME/KEY: misc_feature
 23 <222> LOCATION: (1)..(270)
 24 <223> OTHER INFORMATION: CDS, function = "inhibitor of barnase", product =
 25 "barstar"
 27 <400> SEQUENCE: 1
 28 atgaaaaaag cagtcattaa cggggaacaa atcagaagta tcagcgacct ccaccagaca 60
 29 ttgaaaaagg agcttgccct tccggaatac tacgggtgaaa acctggacgc tttatgggat 120
 30 tgtctgaccg gatgggtgga gtaccgctc gttttggaat ggaggcagtt tgaacaaagc 180
 31 aagcagctga ctgaaaatgg cgccgagagt gtgcttcagg ttttccgtga agcgaaagcg 240
 32 gaaggctgcg acatcaccat catactttct 270
 34 <210> SEQ ID NO: 2
 35 <211> LENGTH: 90
 36 <212> TYPE: PRT
 37 <213> ORGANISM: Artificial Sequence ✓
 39 <220> FEATURE:
 40 <223> OTHER INFORMATION: Description of Artificial Sequence: protein ✓
 42 <400> SEQUENCE: 2
 43 Met Lys Lys Ala Val Ile Asn Gly Glu Gln Ile Arg Ser Ile Ser Asp
 44 1 5 10 15
 46 Leu His Gln Thr Leu Lys Lys Glu Leu Ala Leu Pro Glu Tyr Tyr Gly
 47 20 25 30
 49 Glu Asn Leu Asp Ala Leu Trp Asp Cys Leu Thr Gly Trp Val Glu Tyr
 50 35 40 45
 52 Pro Leu Val Leu Glu Trp Arg Gln Phe Glu Gln Ser Lys Gln Leu Thr
 53 50 55 60
 55 Glu Asn Gly Ala Glu Ser Val Leu Gln Val Phe Arg Glu Ala Lys Ala
 56 65 70 75 80
 58 Glu Gly Cys Asp Ile Thr Ile Ile Leu Ser
 59 85 90
 62 <210> SEQ ID NO: 3
 63 <211> LENGTH: 273
 64 <212> TYPE: DNA
 65 <213> ORGANISM: Artificial Sequence ✓
 67 <220> FEATURE:
 68 <223> OTHER INFORMATION: Description of Artificial Sequence: other nucleic ✓

RAW SEQUENCE LISTING

DATE: 10/19/2001

PATENT APPLICATION: US/09/970,921

TIME: 10:37:39

Input Set : A:\2428-0108P.txt

Output Set: N:\CRF3\10192001\I970921.raw

```

69      acid, "improved barstar DNA"
71 <220> FEATURE:
72 <221> NAME/KEY: misc_feature
73 <222> LOCATION: (1)..(273)
74 <223> OTHER INFORMATION: CDS
76 <400> SEQUENCE: 3
77 atggccaaga aggctgtcat caacggggag cagatcagga gcatcagcga cctccaccag 60
78 accctcaaga aggagcttgc ccttccggag tactacggtg agaacctcga cgccctctgg 120
79 gactgcctca ccggctgggt ggagtacccc ctctgtgttg agtggaggca gttcgagcag 180
80 agcaagcagc tcaccgagaa cggcgccgag agcgtgctcc aagtgttcag ggaggccaag 240
81 gccgagggct gcgacatcac catcatcctc tcc 273
83 <210> SEQ ID NO: 4
84 <211> LENGTH: 91
85 <212> TYPE: PRT
86 <213> ORGANISM: Artificial Sequence
88 <220> FEATURE:
89 <223> OTHER INFORMATION: Description of Artificial Sequence: protein
91 <400> SEQUENCE: 4
92 Met Ala Lys Lys Ala Val Ile Asn Gly Glu Gln Ile Arg Ser Ile Ser
93   1             5             10             15
95 Asp Leu His Gln Thr Leu Lys Lys Glu Leu Ala Leu Pro Glu Tyr Tyr
96             20             25             30
98 Gly Glu Asn Leu Asp Ala Leu Trp Asp Cys Leu Thr Gly Trp Val Glu
99             35             40             45
101 Tyr Pro Leu Val Leu Glu Trp Arg Gln Phe Glu Gln Ser Lys Gln Leu
102             50             55             60
104 Thr Glu Asn Gly Ala Glu Ser Val Leu Gln Val Phe Arg Glu Ala Lys
105 65             70             75             80
107 Ala Glu Gly Cys Asp Ile Thr Ile Ile Leu Ser
108             85             90
111 <210> SEQ ID NO: 5
112 <211> LENGTH: 4032
113 <212> TYPE: DNA
114 <213> ORGANISM: Artificial Sequence
116 <220> FEATURE:
117 <223> OTHER INFORMATION: Description of Artificial Sequence: other nucleic
118      acid, "plasmid pMV71"
120 <220> FEATURE:
121 <221> NAME/KEY: misc_feature
122 <222> LOCATION: (1999)..(3400)
123 <223> OTHER INFORMATION: label = PRAC1, "promoter region of rice actin gene
124      - contains an intron in the leader"
126 <220> FEATURE:
127 <221> NAME/KEY: misc_feature
128 <222> LOCATION: (3401)..(3676)
129 <223> OTHER INFORMATION: label = barstar, "barstar DNA"
131 <220> FEATURE:
132 <221> NAME/KEY: misc_feature
133 <222> LOCATION: (3677)..(4003)

```

RAW SEQUENCE LISTING

DATE: 10/19/2001

PATENT APPLICATION: US/09/970,921

TIME: 10:37:39

Input Set : A:\2428-0108P.txt

Output Set: N:\CRF3\10192001\I970921.raw

```

134 <223> OTHER INFORMATION: label = 3'nos, "region containing 3' untranslated
135     end of nopaline synthase gene of Agrobacterium
136     T-DNA"
138 <220> FEATURE:
139 <221> NAME/KEY: misc_feature
140 <222> LOCATION: (3399)..(3404)
141 <223> OTHER INFORMATION: label = NcoI, "NcoI recognition site"
143 <220> FEATURE:
144 <221> NAME/KEY: misc_feature
145 <222> LOCATION: (4016)..(4021)
146 <223> OTHER INFORMATION: label = KpnI, "KpnI recognition site"
148 <400> SEQUENCE: 5
149 caagcttgac gtcaggtggc acttttcggg gaaatgtgcg cggaaccctt atttgtttat 60
150 ttttctaaat acattcaaat atgtatccgc tcatgagaca ataaccctga taaatgcttc 120
151 aataatattg aaaaaggaag agtatgagta ttcaacattt ccgtgtcgcc cttattccct 180
152 tttttgcggc attttgcctt cctgtttttg ctcaccaga aacgctggtg aaagtaaaaag 240
153 atgctgaaga tcagttgggt gcacgagtg gttacatcga actggatctc aacagcggtg 300
154 agatccttga gagttttcgc cccgaagaac gttttccaat gatgagcact tttaaagttc 360
155 tgctatgtgg cgcggtatta tcccgattg acgcccggca agagcaactc ggtcgccgca 420
156 tacactattc tcagaatgac ttggttgagt actcaccagt cacagaaaag catcttacgg 480
157 atggcatgac agtaagagaa ttatgcagtg ctgccataac catgagtgat aacactgcgg 540
158 ccaacttact tctgacaacg atcggaggac cgaaggagct aaccgctttt ttgcacaaca 600
159 tgggggatca tgtaactcgc cttgatcggt gggaaccgga gctgaatgaa gccataccaa 660
160 acgacgagcg tgacaacacg atgcctgtag caatggcaac aacgttgcg aaactattaa 720
161 ctggcgaaact acttactcta gcttcccggc aacaattaat agactggatg gaggcggata 780
162 aagttgcagg accacttctg cgctcggccc ttccggctgg ctggtttatt gctgataaat 840
163 ctggagccgg tgagcgtggg tctcgcggtg tcattgcagc actggggcca gatggtgaagc 900
164 cctcccgat cgtagttatc tacacgacgg ggagtcaggc aactatggat gaacgaaata 960
165 gacagatcgc tgagataggt gcctcactga ttaagcattg gtaactgtca gaccaagttt 1020
166 actcatatat acttttagatt gattttaaac ttcatTTTTT attttaaagg atctaggtga 1080
167 agatcctttt tggctcgagt ctcatgacca aaatccctta acgtgagttt tctgtccact 1140
168 gagcgtcaga ccccgtagaa aagatcaaag gatcttcttg agatcctttt tttctgcgcg 1200
169 taatctgctg cttgcaaaca aaaaaaccac cgctaccagc ggtggtttgt ttgccggatc 1260
170 aagagctacc aactcttttt ccgaaggtaa ctggcttcag cagagcgagc ataccaaata 1320
171 ctgtccttct agtgtagccg tagttaggcc accacttcaa gaactctgta gcaccgccta 1380
172 catacctcgc tctgctaate ctgttaccag tggctgctgc cagtggcgat aagtcgtgtc 1440
173 ttaccggggt ggactcaaga cgatagttac cggataaagg gcagcggtcg ggctgaacgg 1500
174 ggggttcgtg cacacagccc agcttggagc gaacgacctt caccgaactg agatacctac 1560
175 agcgtgagca ttgagaaaag gccacgcttc ccgaaggagg aaaggcggac aggtatccgg 1620
176 taagcggcag ggtcggaaca ggagagcgca cgaggggagc tccaggggga aacgcctggt 1680
177 atctttatag tctgtcgggg ttccgccacc tctgacttga gcgtcgattt ttgtgatgct 1740
178 cgtcaggggg gcggagccta tggaaaaacg ccagcaacgc ggccttttta cggttcctgg 1800
179 ccttttgcgt gccttttgc ccatgttct ttctgctgt atccccgat tctgtggata 1860
180 accgtattac cgcctttgag tgagctgata ccgtcgcggc cagccgaacg accgagcgca 1920
181 gcgagtcagt gagcgaggaa gcggaagagc gcccaatacg caaaccgcct ctccccgcgc 1980
182 gttggcctga tcagaatttc gaggtcattc atatgcttga gaagagagtc gggatagtc 2040
183 aaaataaaa aaaggtgaaga ttacctggtc aaaagtgaac acatcagtta aaaggtggtg 2100
184 taaagtaaaa tctcggtaat aaaaggtggc ccaaagtga atttactctt ttctactatt 2160
185 ataaaaattg aggatgtttt tgtcgggtact ttgatacgtc atttttgtat gaattggttt 2220

```

RAW SEQUENCE LISTING

DATE: 10/19/2001

PATENT APPLICATION: US/09/970,921

TIME: 10:37:39

Input Set : A:\2428-0108P.txt

Output Set: N:\CRF3\10192001\I970921.raw

```

186 ttaagtttat tcgcttttgg aaatgcatat ctgtatttga gtcgggtttt aagttcgttt 2280
187 gctttttgtaa atacagaggg atttgtataa gaaatatctt taaaaaaacc catatgctaa 2340
188 tttgacataa tttttgagaa aaatatatat tcaggcggaat tctcacaatg aacaataata 2400
189 agattaaaaat agctttcccc cgttgcagcg catgggtatt ttttctagta aaaataaaag 2460
190 ataaacttag actcaaaaca tttacaaaaa caaccacctaa agttcctaaa gcccaaagt 2520
191 ctatccacga tccatagcaa gcccagccca acccaaccca acccaaccca ccccgagcca 2580
192 gccaaactgga caatagtctc cacacccccc cactatcacc gtgagttgtc cgcacgcacc 2640
193 gcacgtctcg cagccaaaaa aaaaaaaaga aagaaaaaaa agaaaaagaa aaaacagcag 2700
194 gtgggtccgg gtcgtggggg ccggaaacgc gaggaggatc gcgagccagc gacgaggccg 2760
195 gccctccctc cgcttccaaa gaaacgcgcc ccctcgccac tatatacata cccccccctc 2820
196 tcctcccatc cccccaaccc taccaccacc accaccacca cctccacctc ctccccctc 2880
197 gctgccggac gacgagctcc tccccctcc cctccgcgcg ccgcccgcgc ggtaaccacc 2940
198 ccgcccctct cctctttctt tctccgtttt ttttttccgt ctccgtctcg atctttggcc 3000
199 ttggtagttt ggggtgggca gaggcggctt cgtgcgcgcc cagatcggtg cgcgggagg 3060
200 gcgggatctc gcggtgggg ctctcgccg cgtggatccg gcccggatct cgcggggaat 3120
201 ggggctctcg gatgtagatc tgcgatccgc cgttggtggg ggagatgatg gggggtttaa 3180
202 aatttccgcc atgctaaaca agatcaggaa gaggggaaa gggcactatg gtttataatt 3240
203 ttatataatt ctgctgcttc gtcaggctta gatgtgctag atctttcttt cttctttttg 3300
204 tgggtagaat ttgaatccct cagcattgtt catcggtagt tttctttttc atgatttg 3360
205 acaaatgcag cctcgtgcgc agcttttttg taggtagacc atggccaaaa aagcagtc 3420
206 taacggggaa caaatcagaa gtatcagcga cctccaccag acattgaaaa aggagcttgc 3480
207 ccttccggaa tactacgggtg aaaacctgga cgctttatgg gattgtctga ccggatgggt 3540
208 ggagtacccg ctcgtttttg aatggaggca gtttgaacaa agcaagcagc tgactgaaaa 3600
209 tggcgccgag agtgtgcttc aggttttccg tgaagcgaaa gcggaaggct gcgacatcac 3660
210 catcatactt tcttaatacg atcaatggga gatgaacaat atggaaacac aaaccgcaa 3720
211 gctagcttgg ctctagagga tccgaagcag atcgttcaaa catttgcaa taaagtttct 3780
212 taagattgaa tctgttgcc ggtcttgcca tgattatcat ataatttctg ttgaattacg 3840
213 ttaagcatgt aataattaac atgtaatgca tgacgttatt tatgagatgg gtttttatga 3900
214 ttagagtccc gcaattatac atttaatacg cgatagaaaa caaaatatag cgcgcaaact 3960
215 aggataaatt atcgcgcgcg gtgtcatcta tgttactaga tcgggaagat ccccggtac 4020
216 cgagctcgaa tt 4032

```

218 <210> SEQ ID NO: 6

219 <211> LENGTH: 563

220 <212> TYPE: DNA

221 <213> ORGANISM: Artificial Sequence ✓

223 <220> FEATURE:

224 <223> OTHER INFORMATION: Description of Artificial Sequence: other nucleic acid, "part of plasmid pLH43" ✓

227 <220> FEATURE:

228 <221> NAME/KEY: misc_feature

229 <222> LOCATION: (1)..(6)

230 <223> OTHER INFORMATION: label = NcoI, "NcoI recognition site"

232 <220> FEATURE:

233 <221> NAME/KEY: misc_feature

234 <222> LOCATION: (3)..(278)

235 <223> OTHER INFORMATION: label = synb*, "improved barstar DNA"

237 <220> FEATURE:

238 <221> NAME/KEY: misc_feature

239 <222> LOCATION: (279)..(545)

RAW SEQUENCE LISTING

DATE: 10/19/2001

PATENT APPLICATION: US/09/970,921

TIME: 10:37:39

Input Set : A:\2428-0108P.txt

Output Set: N:\CRF3\10192001\I970921.raw

```

240 <223> OTHER INFORMATION: label = 3'nos, "region containing 3' untranslated
241     end of nopaline synthase gene of Agrobacterium
242     T-DNA"
244 <220> FEATURE:
245 <221> NAME/KEY: misc_feature
246 <222> LOCATION: (558)..(563)
247 <223> OTHER INFORMATION: label = KpnI, "KpnI recognition site"
249 <400> SEQUENCE: 6
250 ccatggccaa gaaggctgtc atcaacgggg agcagatcag gagcatcagc gacctccacc 60
251 agacctcaa gaaggagctt gcccttcgag agtactacgg tgagaacctc gacgccctct 120
252 gggactgcct caccggctgg gtggagtacc ccctcgtgtt ggagtggagg cagttcgagc 180
253 agagcaagca gtcaccgag aacggcgccg agagcgtgct ccaagtgttc agggaggcca 240
254 aggccgaggg ctgcgacatc accatcatcc tctcctgatg gatccgaagc agatcgttca 300
255 aacatttggc aataaagttt cttaagattg aatcctgttg ccggtcttgc gatgattatc 360
256 atataatttc tgttgaatta cgtaagcat gtaataatta acatgtaatg catgacgtta 420
257 tttatgagat gggtttttat gattagagtc ccgcaattat acatttaata cgcgatagaa 480
258 aacaaaatat agcgcgcaaa ctaggataaa ttatcgcgcg cgggtgtcatc tatgttacta 540
259 gatcgggaag atccccgggt acc                                     563
261 <210> SEQ ID NO: 7
262 <211> LENGTH: 5349
263 <212> TYPE: DNA
264 <213> ORGANISM: Artificial Sequence ✓
266 <220> FEATURE:
267 <223> OTHER INFORMATION: Description of Artificial Sequence: other nucleic ✓
268     acid, "T-DNA of pTTS243"
270 <220> FEATURE:
271 <221> NAME/KEY: misc_feature
272 <222> LOCATION: Complement((1)..(25))
273 <223> OTHER INFORMATION: label = RB, "T-DNA right border"
275 <220> FEATURE:
276 <221> NAME/KEY: misc_feature
277 <222> LOCATION: Complement((98)..(331))
278 <223> OTHER INFORMATION: label = 3'g7, "region containing 3' untranslated
279     end of Agrobacterium T-DNA gene 7"
281 <220> FEATURE:
282 <221> NAME/KEY: misc_feature
283 <222> LOCATION: Complement((332)..(883))
284 <223> OTHER INFORMATION: label = bar, "region coding for phosphinthricin
285     acetyl transferase"
287 <220> FEATURE:
288 <221> NAME/KEY: misc_feature
289 <222> LOCATION: Complement((884)..(2258))
290 <223> OTHER INFORMATION: label = P35S, "35S promoter of Cauliflower Mosaic
291     Virus"
293 <220> FEATURE:
294 <221> NAME/KEY: misc_feature
295 <222> LOCATION: (2281)..(3969)
296 <223> OTHER INFORMATION: label = PE1, "promoter of E1 gene of rice (WO
297     92/13956)"

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/970,921

DATE: 10/19/2001

TIME: 10:37:40

Input Set : A:\2428-0108P.txt

Output Set: N:\CRF3\10192001\I970921.raw

L:9 M:270 C: Current Application Number differs, Replaced Current Application Number